

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 22 MARS 2005

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

DOCUMENT DE
PRIORITÉ
PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA RÈGLE
17.1. a) OU b)

Martine PLANCHE

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint-Petersbourg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr





BREVET D'INVENTION

CERTIFICAT D'UTILITE

26bis, rue de Saint-Petersbourg
75800 Paris Cédex 08
Téléphone: 01 53.04.53.04 Télécopie: 01.42.94.86.54

Code de la propriété intellectuelle-livre VI

REQUÊTE EN DÉLIVRANCE

DATE DE REMISE DES PIÈCES: N° D'ENREGISTREMENT NATIONAL: DÉPARTEMENT DE DÉPÔT: DATE DE DÉPÔT:	Brigitte RUELLAN-LEMONNIER THOMSON 46, quai Alphonse Le Gallo 92648 BOULOGNE cedex France
Vos références pour ce dossier: PF040032	

1 NATURE DE LA DEMANDE	
Demande de brevet	
2 TITRE DE L'INVENTION	
PROCÉDE DE DECODAGE D'UNE SEQUENCE D'IMAGES CODÉE AVEC ECHELONNABILITE SPATIALE ET TEMPORELLE	
3 DECLARATION DE PRIORITE OU REQUETE DU BENEFICE DE LA DATE DE DEPOT D'UNE DEMANDE ANTERIEURE FRANCAISE	Pays ou organisation Date N°
4-1 DEMANDEUR	
Nom Suivi par Rue Code postal et ville Pays Nationalité Forme juridique N° SIREN Code APE-NAF N° de téléphone N° de télécopie	THOMSON LICENSING S.A. RUELLAN Brigitte 46, quai Alphonse Le Gallo 92100 Boulogne France France Société anonyme 383 461 191 322A + 33 1 41 86 50 00 + 33 1 41 86 56 33
5A MANDATAIRE	
Nom Prénom Qualité Cabinet ou Société Rue Code postal et ville N° de téléphone N° de télécopie Courrier électronique	RUELLAN-LEMONNIER Brigitte Mandataire professionnel communautaire, Pouvoir général: 12778 THOMSON 46, quai Alphonse Le Gallo 92648 BOULOGNE cedex + 33 1 41 86 52 77 + 33 1 41 86 56 33 brigitte.ruellan@thomson.net

6 DOCUMENTS ET FICHIERS JOINTS		Fichier électronique	Pages	Détails	
Texte du brevet		textebrevet.pdf	16	D 13, R 2, AB 1	
Dessins		dessins.pdf	4	page 4, figures 7	
Désignation d'inventeurs					
Pouvoir général					
7 MODE DE PAIEMENT					
Mode de paiement		Prélèvement du compte courant			
Numéro du compte client		3334			
8 RAPPORT DE RECHERCHE					
Etablissement immédiat					
9 REDEVANCES JOINTES		Devise	Taux	Quantité	Montant à payer
062 Dépôt		EURO	0.00	1.00	0.00
063 Rapport de recherche (R.R.)		EURO	320.00	1.00	320.00
Total à acquitter		EURO			320.00

La loi n°78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

Signé par

Signataire: FR, THOMSON, B.Ruellan

Emetteur du certificat: DE, D-Trust GmbH, D-Trust for EPO 2.0

Fonction

Mandataire agréé (Mandataire 1)



BREVET D'INVENTION CERTIFICAT D'UTILITÉ

Réception électronique d'une soumission

Il est certifié par la présente qu'une demande de brevet (ou de certificat d'utilité) a été reçue par le biais du dépôt électronique sécurisé de l'INPI. Après réception, un numéro d'enregistrement et une date de réception ont été attribués automatiquement.

Demande de brevet : X

Demande de CU :

DATE DE RECEPTION	2 mars 2004	
TYPE DE DEPOT	INPI (PARIS) - Dépôt électronique	Dépôt en ligne: X
N° D'ENREGISTREMENT NATIONAL ATTRIBUE PAR L'INPI	0450419	Dépôt sur support CD:
Vos références pour ce dossier	PF040032	

DEMANDEUR

Nom ou dénomination sociale	THOMSON LICENSING S.A.
Nombre de demandeur(s)	1
Pays	FR

TITRE DE L'INVENTION

PROCEDE DE DECODAGE D'UNE SEQUENCE D'IMAGES CODEE AVEC ECHELONNABILITE SPATIALE ET TEMPORELLE

DOCUMENTS ENVOYES

package-data.xml	Requetefr.PDF	application-body.xml
Design.PDF	ValidLog.PDF	fee-sheet.xml
FR-office-specific-info.xml	Comment.PDF	textebrevet.pdf
dessins.pdf	indication-bio-deposit.xml	request.xml

EFFECTUE PAR

Effectué par:	B.Ruellan
Date et heure de réception électronique:	2 mars 2004 17:25:47
Empreinte officielle du dépôt	0D:F4:32:4C:1F:95:D8:20:E8:1F:D2:CC:0D:79:2F:0E:EA:4D:D7:42

/ INPI PARIS, Section Dépôt /

SIEGE SOCIAL
INSTITUT 26 bis, rue de Saint Petersburg
NATIONAL DE 75000 PARIS cedex 08
LA PROPRIETE Téléphone : 01 53 04 53 04
INDUSTRIELLE Télécopie : 01 42 93 59 30

5 L'invention concerne un procédé de codage et de décodage vidéo d'une séquence d'image codée avec échelonnabilité spatiale et temporelle, par analyse temporelle hiérarchique exploitant le filtrage temporel compensé en mouvement.

10 Le domaine est celui de la compression vidéo basée sur des schémas à échelonnabilité spatiale et/ou temporelle dits aussi « scalables ». Il s'agit par exemple d'un codage en ondelettes 2D+t comportant un filtrage temporel compensé en mouvement.

Une chaîne de codage-extraction-décodage échelonnable est illustrée sur la figure 1.

15 Les images source sont transmises à un circuit de codage vidéo scalable 1. Le train binaire originel obtenu est traité par un extracteur 2 pour donner un train binaire extrait. Ce dernier est décodé par le circuit de décodage 3 qui fournit en sortie la vidéo décodée.

20 L'échelonnabilité ou « scalabilité » permet de générer un train binaire originel duquel on peut extraire des sous trains binaires adaptés à des ensembles donnés, débit, résolution spatiale, fréquence temporelle... Par exemple, si le train binaire scalable originel a été généré à partir d'une séquence vidéo à 25Hz, de résolution 720*480 pixels et sans contrainte de débit, il est possible, après avoir extrait de ce train binaire les données
25 adéquates, d'obtenir un sous-train binaire, par exemple de paramètres 1 Mb/s, 12,5Hz, 360*240 pixels, lui-même scalable. Le décodage de ce sous-train binaire extrait génère une vidéo à 12,5Hz, de taille 360*240 pixels.

30 Dans les approches existantes de compression vidéo scalables, le codage et le décodage procèdent de façon identique, sans prise en compte de conditions de fonctionnement telles que le niveau de décomposition temporelle, le débit, la résolution spatiale de la vidéo décodée... En particulier, si le décodage implique de la compensation de mouvement entre images, cette compensation est appliquée de façon identique, sans tenir compte de la taille des images, ni du débit de la vidéo à décoder. Cela résulte en une qualité
35 d'images dégradée, en particulier lorsque la résolution d'images devient petite par rapport à la taille des filtres d'interpolation utilisés pour la compensation de mouvement.

L'invention a pour but de pallier les inconvénients précités.

Un des objets de l'invention est un procédé de décodage d'une séquence d'images codée avec échelonnabilité spatiale et temporelle, les données codées comportant des informations de mouvement, comprenant une
5 étape de synthèse temporelle hiérarchique effectuant un filtrage temporel compensé en mouvement, ou MCTF, d'images d'un niveau de décomposition en fréquence à partir des dites informations de mouvement, pour fournir des images d'un niveau de décomposition inférieur, caractérisé en ce que, lors
10 d'une opération de filtrage temporel compensé en mouvement, la résolution choisie pour l'exploitation des informations de mouvement et la complexité des filtres d'interpolation utilisés sont fonction d'un scénario de décodage, à savoir des résolutions spatiale, temporelle et du débit sélectionnés pour le décodage ou bien du niveau de décomposition temporelle correspondant ou d'une combinaison de ces paramètres.

15 Selon une mise en œuvre particulière, le nombre de coefficients du filtre d'interpolation utilisé pour la compensation en mouvement dépend du scénario de décodage ou du niveau de décomposition temporelle.

Selon une mise en œuvre particulière, l'étape de synthèse temporelle hiérarchique est un décodage de coefficients ondelettes avec filtrage
20 compensé en mouvement.

L'invention concerne également un procédé de codage d'une séquence d'images de résolution spatiale donnée, avec échelonnabilité spatiale et temporelle, comprenant une étape d'analyse temporelle hiérarchique effectuant un filtrage temporel compensé en mouvement, ou MCTF, d'images
25 d'un niveau de décomposition en fréquence, à partir d'informations de mouvement entre ces images, pour donner des images d'un niveau de décomposition supérieur, caractérisé en ce que, lors d'une opération de filtrage temporel compensé en mouvement, la résolution choisie pour l'exploitation des dites informations de mouvement et la complexité des filtres d'interpolation
30 utilisés sont fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

Le procédé, selon une mise en œuvre particulière, comporte une étape d'estimation de mouvement calculée entre deux images d'un niveau de décomposition donné pour réaliser la compensation de mouvement et en ce
35 que la précision de calcul de l'estimation de mouvement est fonction du niveau de décomposition temporelle ou de la dite résolution spatiale des images sources.

L'étape d'analyse temporelle hiérarchique est par exemple un codage en ondelettes avec filtrage compensé en mouvement.

5 L'invention concerne également un décodeur pour la mise en œuvre du procédé de décodage précédemment décrit, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement pour déterminer la résolution du mouvement et le filtre d'interpolation à utiliser dans la compensation en mouvement pour le filtrage compensé en mouvement, en fonction du scénario de décodage, à savoir des résolutions spatiale et temporelles et du débit sélectionnés pour le décodage ou du niveau de décomposition temporelle correspondant ou d'une combinaison de ces paramètres.

10 L'invention concerne également un codeur pour la mise en œuvre du procédé de codage précédemment décrit, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement pour déterminer le filtre d'interpolation à utiliser par le circuit d'analyse temporelle pour la compensation de mouvement en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

15 Selon un mode de réalisation particulier, le codeur est caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement pour déterminer la précision du mouvement calculé par le circuit d'estimation de mouvement, en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

20 La précision du mouvement et les filtres d'interpolation utilisés pour la compensation de mouvement dans le processus de codage et de décodage sont adaptés en fonction de différents paramètres, tels que le niveau de décomposition temporelle sur lequel on procède. Ces filtres sont adaptés, pour le décodage, au débit du flux décodé, à la résolution spatiale ou temporelle de la vidéo décodée.

25 Grâce à cette compensation en mouvement adaptative, la qualité des images est améliorée, la complexité des traitements est réduite.

D'autres particularités et avantages de l'invention ressortiront mieux de la description suivante donnée à titre d'exemple non limitatif et faite en regard des figures annexées qui représentent :

- 35
- la figure 1, une chaîne de codage selon l'art antérieur,
 - la figure 2, un schéma de codage simplifié,
 - la figure 3, un filtrage temporel d'un GOP,

- la figure 4, un filtrage temporel sur deux images,
- la figure 5, un circuit de décodage,
- la figure 6, un organigramme de choix de configuration de mouvement,
- 5 - la figure 7, un deuxième organigramme de choix de configuration de mouvement.

On considère un schéma de codage/décodage basé ondelettes 2D+t, opérant une analyse/synthèse ondelettes le long des trajectoires de mouvement. Le système opère sur des groupes d'images ou GOPs.

10 L'architecture globale du codeur est décrite en figure 2.

Les images sources sont transmises à un circuit d'analyse temporelle 4 qui effectue une analyse temporelle compensée en mouvement ou MCTF, acronyme de l'anglais motion compensated temporal filtering, pour
15 obtenir les différentes bandes temporelles fréquentielles. Les images sont transmises à un circuit d'estimation de mouvement 7 qui calcule les champs de mouvement. Ces champs sont transmis à un circuit de « pruning » 10 qui effectue un « élagage » ou une simplification des informations de mouvement calculées par le circuit d'estimation de mouvement pour contrôler le coût du
20 mouvement. Les champs de mouvement ainsi simplifiés sont transmis au circuit d'analyse temporelle afin de définir les filtres d'analyse. Ils sont aussi transmis à un circuit de codage 11 qui effectue le codage des champs de mouvement simplifiés.

Les images résultantes de l'analyse temporelle sont transmises à
25 un circuit d'analyse spatiale 5 qui réalise un codage sous-bande de l'image bande basse et des images bandes hautes obtenues par l'analyse temporelle. Les coefficients ondelettes spatio-temporels ainsi obtenus sont finalement codés par un codeur entropique 6. Ce codeur fournit en sortie un ensemble de paquets binaires correspondant aux couches de scalabilités superposées, à la
30 fois en qualité, en résolutions spatiales et temporelles. Un paquetiseur 12 effectue la fusion de ces paquets binaires avec les données de mouvement provenant du circuit de codage 11 pour fournir le train binaire échelonné final.

Les images aux différents niveaux de décomposition temporelle
35 sont transmises par le circuit d'analyse temporelle 4 vers le circuit d'estimation de mouvement 7 comportant un premier circuit de choix de configuration de mouvement. Ce circuit, non représenté sur la figure, définit les conditions de

fonctionnement du circuit d'estimation de mouvement en fonction des différents niveaux de décomposition des images. De manière optionnelle, les informations de mouvement, une fois simplifiées via le circuit de pruning 10, sont transmises au circuit d'analyse temporelle à travers un circuit de commutation de mode 9.

5 Ce circuit permet de tester la qualité de l'estimation de mouvement en testant par exemple le nombre de pixels connectés entre l'image courante et l'image précédente, à un niveau de décomposition donné, et peut imposer au circuit d'analyse temporelle un codage en mode intra ou un codage en mode prédictif, c'est à dire un filtrage de l'image courante avec l'image suivante et non

10 précédente, lorsque cette qualité de mouvement est insuffisante. Le choix entre le mode intra et prédictif dépend par exemple de la qualité de l'estimation de mouvement entre l'image courante et l'image suivante. Le circuit d'analyse temporelle comporte un deuxième circuit de choix de configuration de mouvement, également non représenté sur la figure, qui détermine, en fonction

15 des niveaux de décomposition des images et/ou de la résolution spatiale des images sources, la configuration à adopter pour la compensation de mouvement utilisée dans cette analyse temporelle.

La figure 3 représente de manière synthétique les opérations de

20 filtrage temporel compensé en mouvement réalisées par le circuit d'analyse temporelle 4, avec une décomposition à 4 niveaux pour des GOPs comportant, dans cet exemple, 16 images représentées en traits gras.

Le mode de filtrage utilisé est appelé « lifting ». Au lieu d'utiliser un filtrage complexe pour le codage en ondelettes, exploitant un filtre linéaire de

25 longueur importante, dans notre exemple le filtrage serait effectué sur un groupe de 16 images, ce procédé de filtrage consiste, de manière connue, à « factoriser » le filtre en exploitant des filtres de longueur limitée, par exemple deux si l'on choisit de filtrer les échantillons deux à deux, ce filtrage étant reconduit pour chaque niveau de décomposition. On considère donc le cas où

30 le filtrage dans le sens du mouvement est effectué sur des paires d'images. Le filtrage basse fréquence et haute fréquence, sur chacune des paires du GOP, produit respectivement, au premier niveau de décomposition temporelle, 8 images de fréquence temporelle basse (t-L) et 8 images de fréquence temporelle haute (t-H).

35 Les images de fréquence temporelle basse sont ensuite décomposées à nouveau selon le même procédé. Le filtrage passe bas de ces images fournit 4 nouvelles images de fréquence temporelle basse t-LL et le

filtrage passe-haut de ces mêmes images fournit 4 images de fréquence temporelle haute t-LH. Le troisième niveau de décomposition fournit 2 images de fréquence temporelle basse t-LLL et 2 images de fréquence temporelle haute t-LLH. Le quatrième et dernier niveau fournit une image de fréquence temporelle basse t-LLLL et une image de fréquence temporelle haute t-LLH.

Cette décomposition temporelle est une décomposition 5 bandes temporelles qui génère donc 1 image t-LLLL, 1 image t-LLH, 2 images t-LLH, 4 images t-LH, et 8 images t-H par GOP de 16 images. Les images t-L, t-LL, t-LLL et bien sûr les images originales sont ignorées pour le codage en aval puisqu'elles sont à l'origine de la décomposition en sous-bandes pour fournir les images décorréliées à chaque niveau. Cette décomposition permet donc une nouvelle répartition de l'énergie en générant une image utile de fréquence temporelle basse t-LLLL, qui représente une moyenne de l'ensemble du GOP et dans laquelle est concentrée l'énergie, et quatre niveaux d'images de fréquence temporelle haute de faible énergie, soit 5 bandes de fréquence. Ce sont ces images qui sont transmises au circuit d'analyse spatiale pour une décomposition spatiale en sous-bandes.

Pour réaliser les filtrages, un champ de mouvement est estimé entre chaque paire d'images à filtrer et cela pour chaque niveau. C'est la fonction de l'estimateur de mouvement 7.

Le filtrage d'une paire d'images source A et B consiste par défaut à générer une image basse fréquence temporelle L et une image haute fréquence temporelle H, selon les équations suivantes :

$$\begin{cases} L = (B + MC(A)) / \sqrt{2} \\ H = (A - MC(B)) / \sqrt{2} \end{cases}$$

où MC(I) correspond à l'image I compensée en mouvement.

La somme concerne le filtrage passe-bas, la différence, le filtrage passe-haut.

La figure 4 est une illustration simplifiée du filtrage temporel des deux images successives A et B, l'image A étant la première image selon l'axe des temps et selon l'ordre d'affichage, donnant une image basse fréquence L et une image haute fréquence H.

L'estimation de mouvement se fait par rapport à une image de référence, de l'image courante vers l'image de référence. On cherche pour chaque pixel de l'image courante, son correspondant, s'il existe, dans l'image

de référence, et on lui affecte le vecteur mouvement correspondant. Le pixel de l'image de référence est alors dit connecté.

L'obtention de l'image L nécessite une compensation en mouvement de l'image A. Cette compensation se fait par estimation de mouvement de l'image B vers l'image A en prenant A comme image de référence, un mouvement et donc un vecteur étant alors affecté à chaque pixel de l'image B. La valeur d'un pixel de L est égale, au facteur de forme près, à la somme de la luminance du pixel correspondant de l'image B et de la luminance du pixel ou sous-pixel de A pointé par le vecteur mouvement attribué au pixel correspondant de l'image B. Une interpolation est nécessaire lorsque ce vecteur ne pointe pas sur un pixel de l'image A. Il s'agit ici de prédiction anticipée à partir d'une image de référence passée et d'un calcul de vecteurs mouvements anticipés ou « forward vectors » en anglais, en se référant à la norme MPEG.

L'obtention de l'image H nécessite une compensation en mouvement de l'image B. Cette compensation se fait par estimation de mouvement de l'image A vers l'image B en prenant B comme image de référence, un mouvement et donc un vecteur étant alors affecté à chaque pixel de l'image A. La valeur d'un pixel de H est égale, au facteur de forme près, à la différence de la luminance du pixel correspondant de l'image A et de la luminance du pixel ou sous-pixel de B pointé par le vecteur mouvement attribué au pixel correspondant de l'image A. Une interpolation est nécessaire lorsque ce vecteur ne pointe pas sur un pixel de l'image B. Il s'agit de prédiction différée à partir d'une image de référence postérieure et d'un calcul de vecteurs mouvements différés ou « backward vectors » en anglais, en se référant à la norme MPEG.

D'une manière pratique, seul un champ de vecteurs mouvement est calculé, de A vers B ou de B vers A. L'autre champ de vecteurs mouvement est déduit du premier, générant des pixels non connectés, c'est à dire non affectés d'un vecteur mouvement et correspondant à des trous dans le champ de vecteurs mouvement inverse.

D'une manière pratique, les images de fréquence basse et haute sont calculées comme suit:

$$\begin{cases} H = \frac{B - MC_{A \leftarrow B}(A)}{\sqrt{2}} \\ L = \sqrt{2} \cdot A + MC_{A \leftarrow B}^{-1}(H) \end{cases}$$

Ce filtrage, équivalent à celui précédemment décrit, consiste à calculer d'abord l'image H. Cette image est obtenue par différence point à point de l'image B et de l'image A compensée en mouvement. Ainsi, on retranche à un pixel de B une certaine valeur, interpolée le cas échéant, pointée par le vecteur déplacement dans A, vecteur mouvement calculé lors de l'estimation de mouvement de l'image B vers l'image A.

L'image L est ensuite déduite à partir de l'image H et non plus de l'image B, par addition de l'image A à l'image H compensée en mouvement inverse. $MC_{A \leftarrow B}^{-1}(H)$ correspond à une « décompensation » en mouvement de l'image (H). Ainsi, on ajoute, à un pixel de A ou plus exactement à une valeur normalisée de la luminance du pixel, une certaine valeur, interpolée le cas échéant, située, dans l'image H, à la base d'un vecteur déplacement de B vers A et pointant le pixel de A.

Un même raisonnement peut être tenu au niveau d'un bloc d'image au lieu d'un pixel.

Le circuit d'estimation de mouvement 7 exploite par exemple un algorithme d'estimation de mouvement par appariement de blocs plus connu sous l'appellation anglaise de « block matching ». Un bloc d'image courant est corrélé aux blocs d'une fenêtre de recherche dans l'image de référence pour déterminer le vecteur de mouvement correspondant à la meilleure corrélation. Cette recherche est effectuée non seulement sur les blocs de la fenêtre de recherche obtenus par déplacements successifs d'un pixel en horizontal et vertical mais également sur des blocs interpolés si la précision recherchée est inférieure au pixel. Cette interpolation consiste à calculer les valeurs de luminance des sous-pixels, pour la génération de blocs d'images obtenus par déplacements successifs d'une valeur inférieure à la distance entre deux pixels. Par exemple pour une précision du quart de pixel, un test de corrélation est effectué tous les quarts de pixels, en horizontal et en vertical. Cette interpolation exploite des filtres appelés filtres d'interpolation de l'estimation de mouvement.

Les images pour lesquelles un filtrage temporel compensé en mouvement est à effectuer sont transmises à l'estimateur de mouvement 7 pour qu'il estime le mouvement entre deux images. Ce circuit comporte un premier circuit de choix de configuration de mouvement qui reçoit, en plus de l'information de niveau de décomposition des images, d'autres informations telles que la résolution spatiale des images sources. Ce circuit décide de la

configuration de mouvement en fonction de ce niveau et/ou de la résolution spatiale. Ainsi, par exemple, la précision dans le calcul des vecteurs mouvement dépend du niveau de décomposition temporelle des images traitées. Cette précision est d'autant plus faible que le niveau de décomposition
5 est grand. Les filtres d'interpolation de l'estimateur de mouvement sont configurés pour être adaptés à cette précision de mouvement. Un exemple de configuration est donné par la suite.

Le circuit d'analyse temporelle 4, comme on l'a indiqué plus haut, réalise des compensations de mouvement pour le filtrage temporel des images.
10 Ces opérations de compensations en mouvement nécessitent des opérations d'interpolation utilisant des filtres d'interpolation, et cela pour chaque niveau de décomposition. Le deuxième circuit de choix de configuration de mouvement, dans ce circuit d'analyse temporelle, qui peut être différent du premier, met en œuvre un algorithme de traitement adaptant la précision du mouvement et la
15 complexité du filtre d'interpolation pour la compensation de mouvement en fonction du niveau de décomposition temporelle des images à compenser en mouvement. Comme pour le premier circuit de choix de configuration de mouvement, ces différentes adaptations ou configurations peuvent également dépendre de la résolution spatiale des images sources traitées.

20 Bien évidemment, un codeur ne comportant qu'un seul de ces circuits de choix de configuration est du domaine de l'invention.

Un décodeur selon l'invention est décrit en figure 5. Le flux binaire reçu par le décodeur est transmis à l'entrée d'un circuit de décodage entropique
25 13 qui réalise les opérations inverses du circuit de codage entropique du codeur. Il décode entre autres les coefficients ondelettes spatio-temporels et, le cas échéant, les modes de codage. Ce flux binaire est transmis en parallèle à l'entrée d'un circuit de décodage du mouvement 14 qui décode les champs de mouvement reçus dans le flux binaire pour les transmettre au circuit de
30 synthèse temporelle. Le circuit de décodage entropique 13 est relié à un circuit de synthèse spatiale 15 qui reconstruit les images correspondant aux différentes sous-bandes temporelles. Les coefficients ondelettes temporels provenant du circuit de synthèse spatiale sont transmis à un circuit de synthèse temporelle 16 qui reconstruit les images de sortie à partir de filtres de synthèse
35 temporelle. Le circuit de synthèse temporelle comporte un circuit de choix de configuration de mouvement, non représenté sur la figure, qui détermine, en fonction des conditions du décodage et/ou des niveaux de décomposition des

images, la configuration à adopter pour la compensation de mouvement utilisée dans cette synthèse temporelle. Le circuit de synthèse temporelle est relié à un circuit de post-traitement 17 dont la sortie est la sortie du décodeur. Il s'agit par exemple de post-filtrage permettant de réduire les artéfacts tels que les effets
5 de bloc.

Dans le cas où le codeur exploite d'autres modes de codages que le mode MCTF, par exemple le mode intra et le mode prédictif, un circuit de commutation de filtre temporel est utilisé pour recevoir ces informations de mode de codage provenant du circuit de décodage entropique 13 et pour les
10 transmettre au circuit de synthèse temporelle 16 qui effectue en conséquence les commutations de filtres.

Le circuit de choix de configuration de mouvement reçoit les informations de débit, de résolution spatiale et temporelle et les niveaux de décomposition temporelle. A partir de ces informations ou d'une de ces
15 informations, il choisit, pour la synthèse temporelle, une configuration de compensation de mouvement. Le circuit de synthèse temporelle adapte le filtre d'interpolation en fonction de cette configuration choisie.

Le débit du flux binaire reçu par le décodeur correspond au train binaire extrait. Le codeur scalable transmet généralement le débit le plus élevé
20 qui est le train binaire original, comme on l'a vu précédemment, et l'extracteur, qui peut être commandé par le décodeur, extrait le train binaire correspondant aux résolutions souhaitées. Les informations de débit reçu sont disponibles au décodeur.

Les informations de résolution spatiale, temporelle et de débit définissent un scénario de décodage. Ce scénario est fonction par exemple de
25 l'afficheur exploité par le décodeur, du débit disponible pour la réception des données. C'est à partir de ces informations et/ou du niveau de décomposition temporelle que le circuit de synthèse temporelle est configuré quant aux filtres d'interpolation.

30

Un exemple d'adaptation de la précision du mouvement et du filtre d'interpolation qui dépend de cette précision est donné ci-dessous, pour les opérations d'estimation de mouvement au codeur ou les opérations de compensation de mouvement au codeur ou décodeur :

35

configuration	précision du mouvement	filtres d'interpolation
1	1/4 de pixel	Bilinéaire
2	1/8 de pixel	¼ pixel par interpolation FIR 8 coefficients, puis 1/8 pixel par interpolation bilinéaire

Le filtre de la configuration 2 est très similaire de celui utilisé dans la norme MPEG-4 partie 10 (référence ITU-T Rec. H.264 ISO/IEC 14496-10 AVC).

5 La figure 6 représente un organigramme de décision mis en œuvre par le circuit de choix de configuration de mouvement appartenant au circuit d'analyse temporelle.

L'étape 20 détermine si la résolution de l'image source fournie au codeur est inférieure à celle du format QCIF, de l'anglais Quarter Common Intermediate Format et correspondant à 176 colonnes, 120 lignes. Dans l'affirmative, l'étape suivante est l'étape 23 qui décide la configuration 1.

Dans la négative, l'étape suivante est l'étape 21 qui vérifie le niveau de décomposition temporelle. Si ce niveau est strictement supérieur à 2, l'étape suivante est l'étape 23, la configuration 1 est choisie. Sinon, l'étape 15 suivante est l'étape 22 qui décide la configuration 2.

La figure 7 représente un organigramme de décision pour le décodeur.

L'étape 24 détermine si la résolution de l'image fournie par le décodeur et correspondant au flux binaire extrait est inférieure à celle du format QCIF, 176 colonnes, 120 lignes. Dans l'affirmative, l'étape suivante est l'étape 26 qui choisit la configuration 1.

Dans la négative, l'étape suivante est l'étape 25 qui vérifie le niveau de décomposition temporelle. Si ce niveau est strictement supérieur à 2, l'étape suivante est l'étape 26, la configuration 1 est utilisée. Sinon, l'étape 25 suivante est l'étape 27. Cette étape 27 détermine si la résolution de l'image à décoder est égale à celle du format SD, de l'anglais Standard Definition, 720 colonnes, 480 lignes et si le débit du flux binaire reçu est inférieur à 1,5 Mb/s. Dans l'affirmative, l'étape suivante est l'étape 26 qui décide la configuration 1.

Dans la négative, l'étape 28 est l'étape suivante. Cette étape 28 détermine si la résolution de l'image à décoder est égale à celle du format CIF, 352 colonnes, 240 lignes et si le débit est inférieur à 700 kbits/s. Dans l'affirmative, l'étape suivante est l'étape 26 qui impose la configuration 1.

Dans la négative, la configuration 2 est imposée aux circuits de filtrage temporel.

Le filtre d'interpolation est par exemple du type FIR 8 coefficients, acronyme de l'anglais Finite Impulse Response. Le filtrage est effectué par convolution, en prenant donc en compte les luminances des 4 pixels précédents et suivants le sous-pixel à calculer.

Pour différentes positions au sous-pixel s à $\frac{1}{4}$, $\frac{1}{2}$, et $\frac{3}{4}$, on peut utiliser trois filtres d'interpolation différents du type précédent. La valeur d'un coefficient n est donné par la formule :

$$f(n+s) = \sum_{m=-4}^4 h(m) \frac{\sin \pi(n+s-m)}{\pi(n+s-m)}, \quad 0 < s < 1.$$

s est la position sous-pixel, $s = \frac{1}{4}$, $\frac{1}{2}$, ou $\frac{3}{4}$, n est le numéro du coefficient et $h(m)$ le filtre d'atténuation ou fenêtre de Hamming.

On peut déduire des filtres FIR, par pondération par une fenêtre de Hamming et troncation de ces filtres pondérés.

Pour $s = \frac{1}{4}$, les coefficients sont :

[-0.0110 0.0452 -0.1437 0.8950 0.2777 -0.0812 0.0233 -0.0053]

Pour $s = \frac{1}{2}$, les coefficients sont :

[-0.0053 0.0233 -0.0812 0.2777 0.8950 -0.1437 0.0452 -0.0110]

Pour $s = \frac{3}{4}$, les coefficients sont :

[-0.0105 0.0465 -0.1525 0.6165 0.6165 -0.1525 0.0465 -0.0105]

Avec ces filtres, on peut interpoler au $\frac{1}{4}$, $\frac{1}{2}$ et $\frac{3}{4}$ de pixel. L'interpolation se fait d'abord selon la dimension horizontale, puis verticale. L'interpolation au $\frac{1}{8}$ de pixel est effectuée ensuite par une interpolation bilinéaire à partir des positions au $\frac{1}{4}$ de pixel.

L'exemple d'adaptation donné ci-dessus au niveau du codeur peut s'appliquer de la même manière au niveau du décodeur.

D'une manière générale, le principe est d'utiliser une précision de mouvement limitée et des filtres d'interpolation simples quand on opère à des qualités d'image limitées, c'est-à-dire à bas débit, sur des images de petite taille et à des niveaux de décomposition temporelle élevés. A l'inverse, quand on traite des images de bonne qualité, résolution spatiale grande, débits élevés,

niveaux de décomposition temporelle bas, on utilise une grande précision de mouvement et des filtres d'interpolation sophistiqués. La justification de ce principe est que lorsque les images à filtrer sont pauvres en contenu fréquentiel ou de résolution limitée, il n'est pas utile d'employer des filtres d'interpolation très évolués ni une précision de mouvement très grande.

Les applications de l'invention concernent les codeurs/décodeurs vidéo dits « scalables » utilisés pour la compression/décompression de données, par exemple dans le domaine de la vidéotéléphonie ou transmission vidéo sur internet.

REVENDECATIONS

1 Procédé de décodage d'une séquence d'images codée avec
5 échelonnabilité spatiale et temporelle, les données codées comportant des informations de mouvement, comprenant une étape de synthèse temporelle hiérarchique (16) effectuant un filtrage temporel compensé en mouvement, ou MCTF, d'images d'un niveau de décomposition en fréquence à partir des dites informations de mouvement, pour fournir des images d'un niveau de
10 décomposition inférieur, caractérisé en ce que, lors d'une opération de filtrage temporel compensé en mouvement, la résolution choisie pour l'exploitation des informations de mouvement et la complexité des filtres d'interpolation utilisés sont fonction d'un scénario de décodage, à savoir des résolutions spatiale, temporelle et du débit sélectionnés pour le décodage ou bien du niveau de
15 décomposition temporelle correspondant ou d'une combinaison de ces paramètres.

2 Procédé selon la revendication 1, caractérisé en ce que le
nombre de coefficients du filtre d'interpolation (16) utilisé pour la compensation
20 en mouvement dépend du scénario de décodage ou du niveau de décomposition temporelle.

3 Procédé selon la revendication 1, caractérisé en ce que l'étape
de synthèse temporelle hiérarchique (16) est un décodage de coefficients
25 ondelettes avec filtrage compensé en mouvement.

4 Procédé de codage d'une séquence d'images de résolution
spatiale donnée, avec échelonnabilité spatiale et temporelle, comprenant une
étape d'analyse temporelle hiérarchique (4) effectuant un filtrage temporel
30 compensé en mouvement, ou MCTF, d'images d'un niveau de décomposition en fréquence, à partir d'informations de mouvement entre ces images (7), pour donner des images d'un niveau de décomposition supérieur, caractérisé en ce que, lors d'une opération de filtrage temporel compensé en mouvement (4), la résolution choisie pour l'exploitation des dites informations de mouvement et la
35 complexité des filtres d'interpolation utilisés (9) sont fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

5 Procédé selon la revendication 4, caractérisé en ce qu'il comporte une étape d'estimation de mouvement (7) calculée entre deux images d'un niveau de décomposition donné pour réaliser la compensation de mouvement (4) et en ce que la précision de calcul de l'estimation de mouvement (7) est fonction du niveau de décomposition temporelle ou de la dite résolution spatiale des images sources.

10 6 Procédé selon la revendication 4, caractérisé en ce que l'étape d'analyse temporelle hiérarchique (4) est un codage en ondelettes avec filtrage compensé en mouvement.

15 7 Décodeur pour la mise en œuvre du procédé selon la revendication 1, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement (16) pour déterminer la résolution du mouvement et le filtre d'interpolation à utiliser dans la compensation en mouvement (16) pour le filtrage compensé en mouvement, en fonction du scénario de décodage, à savoir des résolutions spatiale et temporelles et du débit sélectionnés pour le décodage ou du niveau de décomposition temporelle correspondant ou d'une
20 combinaison de ces paramètres

25 8 Codeur pour la mise en œuvre du procédé selon la revendication 4, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement (4) pour déterminer le filtre d'interpolation à utiliser par le circuit d'analyse temporelle pour la compensation de mouvement (4) en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

30 9 Codeur pour la mise en œuvre du procédé selon la revendication 4, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement (7) pour déterminer la précision du mouvement calculé par le circuit d'estimation de mouvement (7), en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

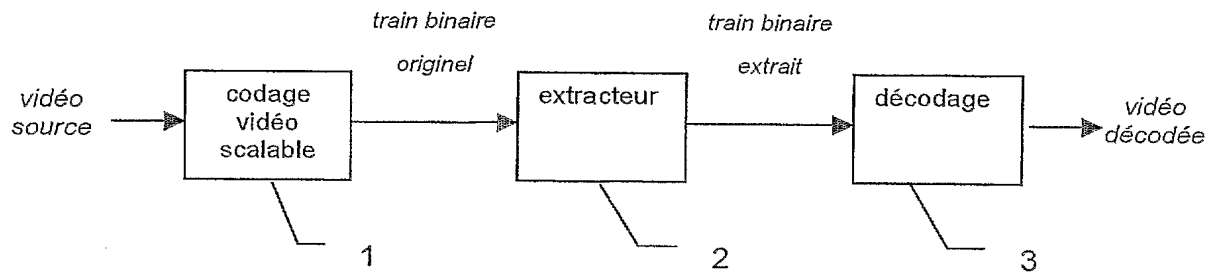


FIG.1

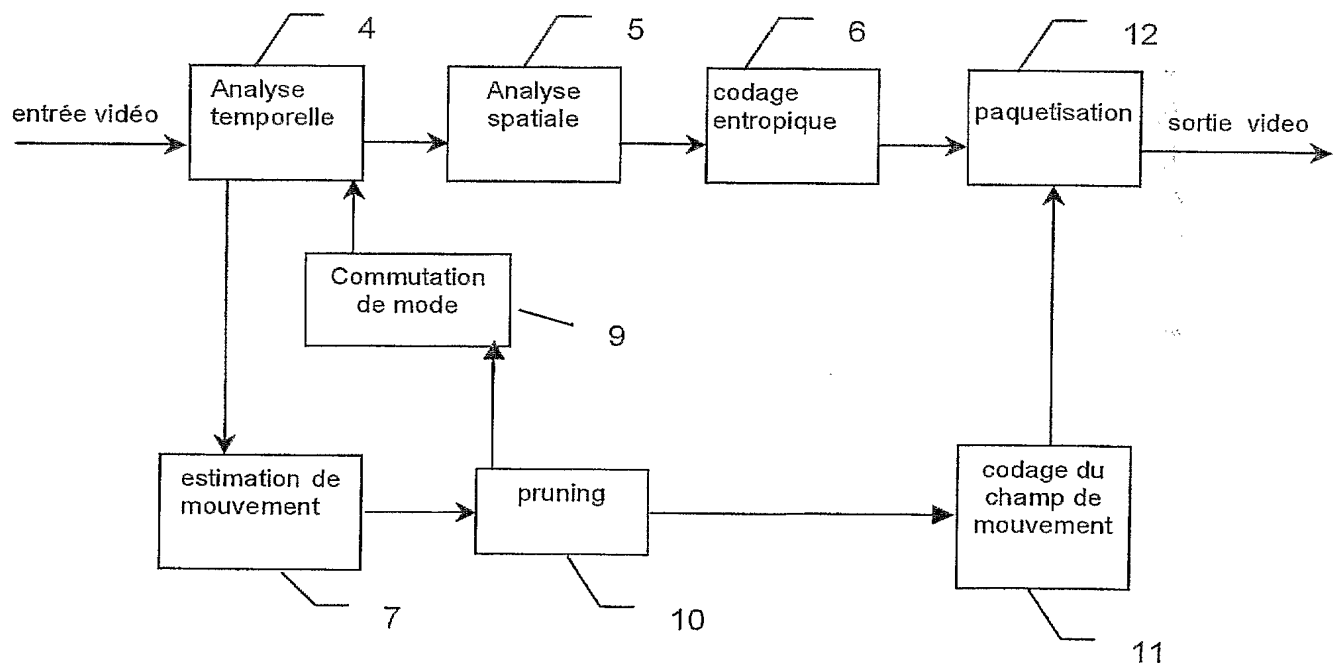


FIG.2

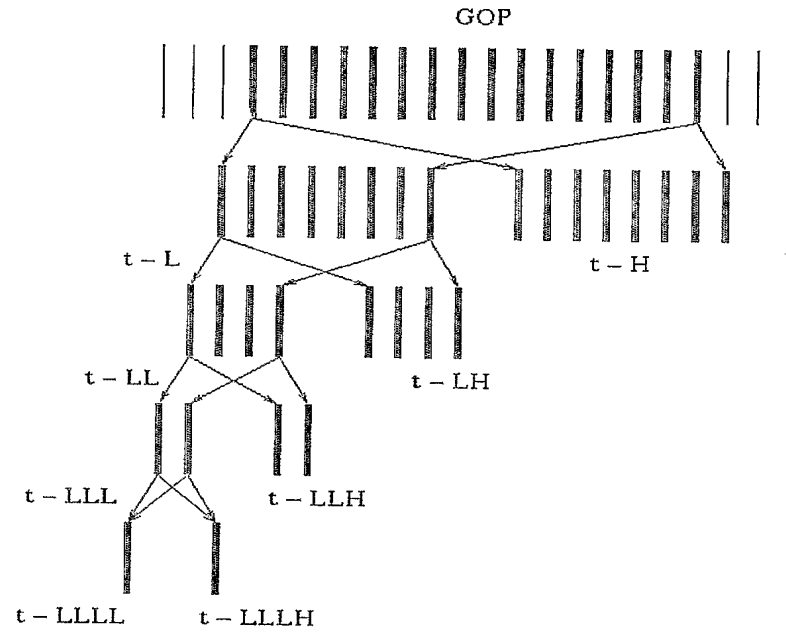


FIG 3

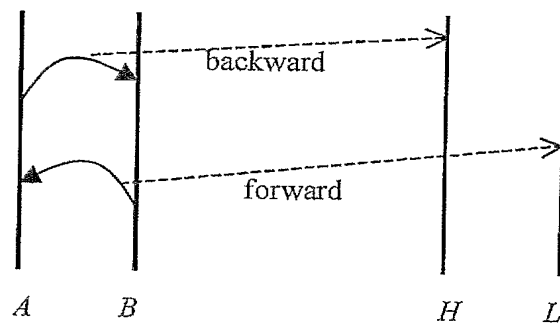


FIG 4

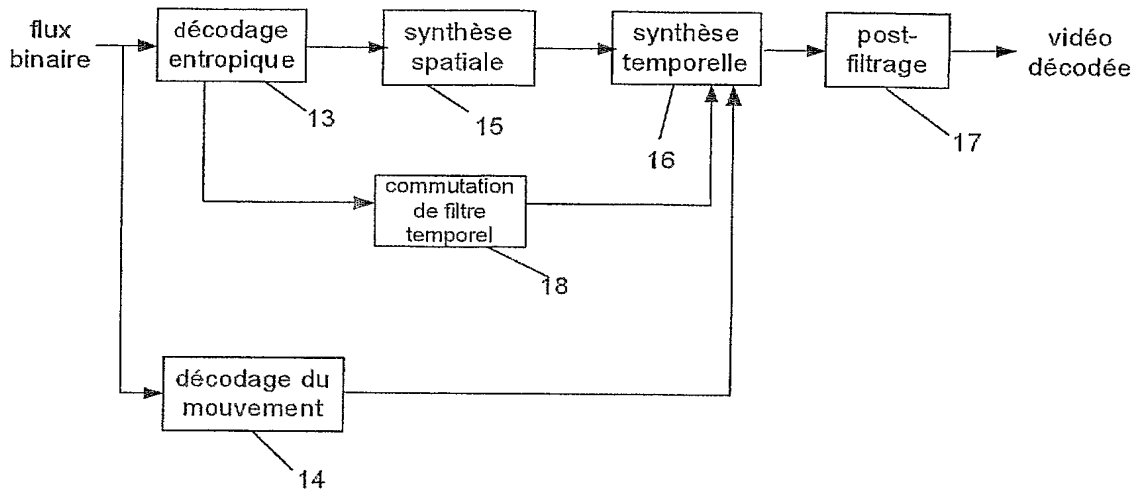


FIG.5

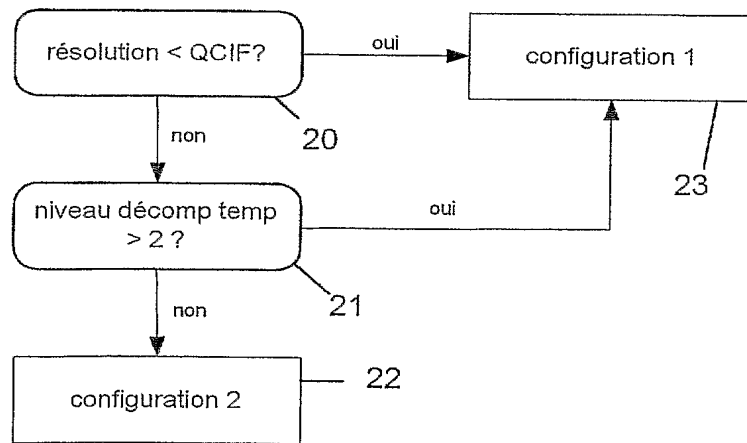


FIG.6

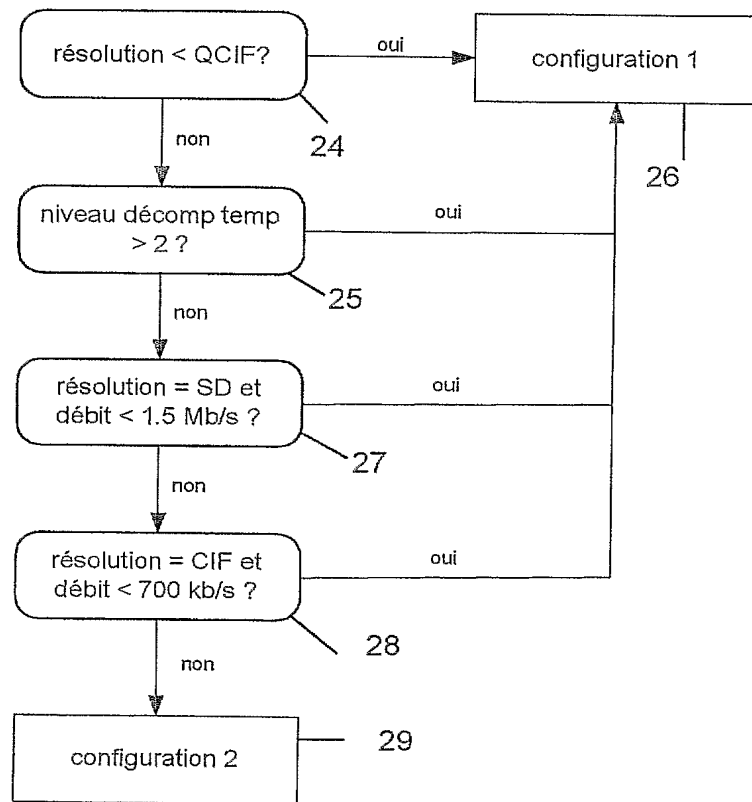


FIG.7



BREVET D'INVENTION CERTIFICAT D'UTILITE

Désignation de l'inventeur

Vos références pour ce dossier	PF040032
N°D'ENREGISTREMENT NATIONAL	
TITRE DE L'INVENTION	
	PROCEDE DE DECODAGE D'UNE SEQUENCE D'IMAGES CODEE AVEC ECHELONNABILITE SPATIALE ET TEMPORELLE
LE(S) DEMANDEUR(S) OU LE(S) MANDATAIRE(S):	
DESIGNE(NT) EN TANT QU'INVENTEUR(S):	
Inventeur 1	
Nom	FRANCOIS
Prénoms	Edouard
Rue	18 Allée du Locar
Code postal et ville	35890 BOURG-DES-COMPTES
Société d'appartenance	THOMSON
Inventeur 2	
Nom	BOISSON
Prénoms	Guillaume
Rue	12 rue Jean Malo-Renault
Code postal et ville	35000 RENNES
Société d'appartenance	THOMSON
Inventeur 3	
Nom	VIERON
Prénoms	Jérôme
Rue	14 Square Louis Massignon
Code postal et ville	35000 RENNES
Société d'appartenance	THOMSON
Inventeur 4	
Nom	ROBERT
Prénoms	Philippe
Rue	7 Allée du Bois Louët
Code postal et ville	35235 THORIGNE-FOUILLARD
Société d'appartenance	THOMSON
Inventeur 5	
Nom	MARQUANT
Prénoms	Gwenaëlle
Rue	1 rue du Général de Gaulle
Code postal et ville	35340 LIFFRE
Société d'appartenance	THOMSON

La loi n°78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

Signé par

Signataire: FR, THOMSON, B.Ruellan

Emetteur du certificat: DE, D-Trust GmbH, D-Trust for EPO 2.0

Fonction

Mandataire agréé (Mandataire 1)

FOT/FR005050108

